Application of Artificial Intelligence in Cyber security

Varun Kharbanda, SP Jain school of global management, Missing Country
Seetharaman A, SP Jain school of global management, Australia
Maddulety K, SP Jain school of global management, Australia

ABSTRACT

Artificial intelligence (AI) has emerged as the most widely applicable field across varied industries. Being an evolving technology, it may be quite useful in sensitive areas such as cyber security where there is a dire need for implementation of AI technologies, such as expert systems, neural networks, intelligent agents, and artificial immune systems. The primary reason for AI fitment to cyber security area is its ability to detect anomalies proactively and predictively in the network, thereby working towards securing the network before the damage related to loss of data and/or reputation is done. There are different types of AI technologies as mentioned above that could be applied in cyber security in its varied forms. In this paper, the emphasis is on specific AI technologies that can bring unique benefits to the cyber security field with its unique applicability to different scenarios. The outcome of this study shows that AI technologies such as expert systems, neural networks, intelligent agents, and artificial immune systems are transforming the landscape for managing cyber threats.

KEYWORDS

Artificial Immune System, Artificial Intelligence, Cyber Risks, Cyber Security, Cyberattacks, Cyberspace, Expert Systems, Intelligent Agent, Machine Learning, Neural Networks

1. INTRODUCTION

Since this article involves understanding how Artificial Intelligence (AI) is going to be applied and its usage in cyber security functions, it is of paramount importance to understand the meaning of AI. AI endeavours to build and recognise smart objects. Employees or users of AI could apply their skills to any industry they deem fit; thus, AI is a broad area in this sense (Russell & Norvig, 2016). At the core, it conveys when the machine starts imitating human intelligence and starts self-learning, leading to unknown solutions – something machine was never capable of doing before. However, algorithmic modelling has made this possible, and AI has become a universal term and function. Moreover, AI technologies such as expert systems, machine learning, deep learning or neural networks, artificial immune system, intelligent agents et cetera are being applied in various fields including but not limited to healthcare, automotive, banking and insurance sectors.

A major problem in corporations these days is how they can guard themselves against possible anomalies. The variety and possibility of these unidentified outbreaks generate the requirements for corporates to prioritise the method in which they protect themselves against such cyberattacks. Thus, each corporate or company is required to understand the attacks that they are most defenceless against.
to reduce the risk of an attack (Fielder, Panaousis, Malacaria, Hankin, & Smeraldi, 2016). With the recent increase in cyberattacks, the uncertainty across the enterprise has increased by leaps and bounds.

Cyberattacks on global company networks, including governments, have kept the cyber-attack prevention teams extremely busy and hence, automation alone will not suffice. However, there is an ever-growing need to handle such cyberattacks proactively and with the ability to predict and resolve them considering the evolution of AI capabilities or techniques, such as neural networks or deep learning, machine learning, expert systems, artificial immune systems, intelligent agents. Such advanced technologies are changing the way cyber security is being managed in modern organisations which are more vulnerable to unknown or unpredictable attacks than ever before. The change is perceived as not just managing these cyberattacks proactively but predicting these attacks and resolving issues before they negatively impact a company’s operations and risk customer’s data.

This article will explore how the cyber security landscape may be transformed after application of emerging and disruptive technologies like AI in handling cyber risks. Most AI tools are used to enhance cyber security, and cyber risk management tools benefit the whole cyber security process, making them more robust to handle security vulnerabilities in organisational networks. Various types of cyber risks, such as ransomware, phishing, data leakage, hacking, trojan horse, computer worm, DOS and DDOS attack, adware and spyware, were usually managed reactively as per traditional practices which involved incident detection and responding accordingly to safeguard company’s network. However, with increasingly sophisticated network attacks, there is a dire need to manage such cyber risks proactively by predicting and protecting companies using AI technologies, such as expert systems, artificial neural network, intelligent agents and artificial immune systems. There are four categories (Early warning/ Prevent, Detect, Reach and Response) of possible scenarios where AI techniques are applied to security issues within the integrated security approach, demonstrating the vast possibilities of the various AI branches (Wirkuttis & Klein, 2017).

2. RESEARCH QUESTIONS

Cyber security or defence management is widespread across the research and corporate sectors. There are several studies regarding the utilisation of AI in cyber security functions throughout industries. The idea here is to address cyber risks through the application of AI as an emerging and/or disruptive technology. The research may encompass a variety of AI technologies, such as artificial expert systems, neural networks, intelligent agents, and artificial immune systems, and how these technologies may change the cyber risk management landscape and their handling. This research addresses the following questions:

1. Does the knowledge acquisition problem affect the application of expert systems in decision-making or problem-solving in cyber security?
2. How do AI technologies such as neural networks help prevent cyberattacks proactively while managing the unknowns?
3. How do intelligent agents fight against cyber assaults such as Distributed Denial of Services (DDoS) and how effective are they?
4. How does an artificial immune system understand changes in patterns and report abnormal behaviour to detect intruders and mitigate risks related to cyber threats or vulnerabilities?

3. RESEARCH OBJECTIVES

In light of the research questions mentioned above, the following research objectives can be derived (also figuratively demonstrated below in Figure 1):
1. To understand causes of the knowledge acquisition problem impact on expert systems in decision-making or problem-solving to improve the management of cyberattacks.
2. To explore how cyberattacks can be prevented proactively using AI neural networks to prevent malicious unknown intrusions.
3. To ascertain the role and capability of Intelligent agents as cyber police to monitor the networks and fight against Distributed Denial of Service (DDoS).
4. To recognise the value of artificial immune systems in understanding the changing patterns and detecting anomalies proactively, leading to mitigation of cyber risks.

4. LITERATURE REVIEW

Various conference papers, journals, books and articles have been searched on databases like ProQuest, EBSCO, Google Scholar, Research gate, eBrary and O’REILLY Safari using keywords such as ‘Artificial Intelligence and Cyber Security’, ‘Artificial Intelligence’ and ‘Cyber-Risks’ to ensure relevancy. Searches were further refined to be the year 2016 onwards. Many articles were reviewed; however, not all are included in this literature review.

4.1 Expert Systems

Expert systems are a widely known rule-based AI tool. These systems are knowledge-driven and were among the first known application of AI. Security professionals widely use expert systems for decision support in cyber environments (Wirkuttis & Klein, 2017). Hence, the primary usage of expert systems is to support decision making and problem-solving tasks and for network intrusion detection across industries such as corporate or business finance, medical or health diagnostics or cyber space. They are used for the simplest of problems linked to diagnostics to the most complex ones involving hybrid systems.

An expert system contains a knowledge base in which proficient and professional facts of knowledge related to a specific application field is kept. It also includes an inference engine and added knowledge regarding the state of affairs to acquire answers. The vacant knowledge base and the extraction engine are together known as the expert system shell; this must be full of information so that it can be utilised. Developing an expert system implies selecting an expert system shell and adaptation, as well as inputting the expert data and knowledge base with accurate data. The latter is much more difficult than the former step as it is more time-consuming (Şeker, 2019).
To spot network anomalies or any foreign intruders, we need a knowledge base, rule sets and further configurations on which expert systems usually run. Various network invasion behaviour related features are saved in the knowledge base and are gathered from the database, which consists of a linked knowledge base and is saved as the web application part. Real-time data packets need to clear the rule sets, and these rule sets are gathered from the database and saved or preserved for the application infrastructure (Anwar & Hassan, 2017).

Rule bases representation is among many different knowledgeable representation forms; however, the success or effective utilisation hinges on the quality of data that is entered in the expert systems' knowledge base. If the input is not of good quality, naturally the output will be impacted as well. Hence, expert systems have a great dependency on the quality of the data or knowledge being used for effective solutions. Therefore, to develop a real-world application for practical uses, the ‘knowledge acquisition problem’ is crucial and must be stressed.

4.2 Neural Network

An intrusion detection system (IDS) is used to notice any illicit foreign activity, attack or intrusion in a network system. Network intrusion detection systems are extensively set-up in current corporate networks. These structures conventionally relied on patterns of notorious outbreaks, but current set-ups comprise of various methods for anomaly detection, threat discovery and classification based on machine learning (Pierazzi, Apruzzese, Colajanni, Guido, & Marchetti, 2017). Artificial Neural Networks (ANNs) can detect and understand these patterns in order to ensure proactive resolution.

The intention to make use of AI in early warning and foreign object invasion discovery is to grow an intelligent assist system to discover threats from the Internet as early as possible on both local area networks (LAN) and wide-area networks (WAN). In this framework, widely utilised Internet protocols like FTP (File Transfer Protocol), SMTP (Simple Mail Transfer Protocol) and HTTP (HyperText Transfer Protocol) should be included, in addition to the latest protocols, for example, SOAP (Simple Object Access Protocol) (Şeker, 2019).

Artificial neural networks (ANNs) mimic the brain neural network of a human being in order to learn, process information and adapt, leading to problem resolution—in essence, they are machine learning algorithms and are classified as a shallow learning (SL) network and deep learning (DL) network. The primary variance amongst the two lies in the number of hidden layers. Shallow learning has set a single hidden layer, whereas deep learning architecture has multiple hidden layers.

The first discriminant is between the conventional machine learning algorithms, referred to as SL versus the latest or more current DL. SL needs a domain or functional expert who can execute the critical task of finding related or accurate data features before performing the SL algorithm. DL depends on a multi-layered illustration of the entered information and can execute selecting features on its own via process defined representation learning (Apruzzese, Colajanni, Ferretti, Guido, & Marchetti, 2018).

ANNs can be effectively utilised to acquire historical network activities and threats to prevent imminent or forthcoming threats from occurring. Compared to conventional techniques used for cyber defence, the great advantage of using ANNs is their learning ability. Patterns that describe normal and abnormal network activities are traditionally defined manually by security professionals based on their expert knowledge. ANNs, however, can be trained to identify such patterns automatically by using previous data that has been transferred over the network (Wirkuttis & Klein, 2017).

4.3 Intelligent Agents

Intelligent agents are software mechanisms that have intelligent behavioural characteristics such as being proactive, communication, understanding the language and responding), making them cut out for the task of effective management to fight distributed denial of service (DDoS) outbreaks. These software mechanisms have planning, changeability, and deep-thinking abilities (Şeker, 2019). DDoS is a type of attack in which a service is denied to an eligible user or becomes inaccessible due to cyberattack.
Intelligent agents help fight against DDoS attacks. To settle legal and additionally corporate or business problems, preliminary cyber-police, which includes intelligent agents (moveable), should be created. Deployment of infrastructure is needed to provide the intelligent agent’s mobility and communication, but should not be accessible for attackers (Anwar & Hassan, 2017).

Developing cyber-police requires the deployment of infrastructure for supporting the automated agents’ excellence and communication; however, it should not be accessible for attackers. This may require support from various Internet service providers. Multi-agent software or tools will provide a wide-ranging operational view of the cyber house, for example, a hybrid multi-agent and neural network-based invasion discovery technique have been anticipated. (Patil, 2016).

### 4.4 Artificial Immune System

Artificial Immune Systems (AISs) are computer-based mathematical models that have sources from biological immune systems which can adapt in a flexible environment and self-learn. Immune systems are responsible for finding unknown invasions or intruders like various types of bacteria, viruses, et cetera and consequently fight against them. AISs are formed to copy natural immune systems generally in the application of cyber safety, and specifically for Intrusion Detection System (IDS) (Kamtam, Kamar, & Patkar, 2016).

The primary purpose of a biological immune system is to arm the human body to fight against unknown molecules called antigens. The immune system has a wonderful recognition system which can detect changes in patterns and report abnormal behaviours in the system. AIS utilises a machine language which instils the functionality of the biological immune system. In contrast to conventional cyber security approaches, AIS principles have an edge due to their ability to detect attacks internally from the network and prevent them from occurring. The advantage of AIS-based IDS is its use of biologically influenced concepts in computation to stop a network attack by determining malicious patterns even before the attack happens (Song, Kim, Tyagi, & Rajasekaran, 2018).

An IDS applied to the principles of an AIS can provide a resolution to the problems that can occur while securing information. This implementation of an IDS differs from the industry standard of implementing machine learning (Cooper, 2017).

<table>
<thead>
<tr>
<th>Type of Agent</th>
<th>Percepts</th>
<th>Actions</th>
<th>Objectives</th>
<th>Environment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthcare diagnosis system</td>
<td>Warning Signs, Outcomes, patient’s response</td>
<td>Queries, tests, a patient’s treatments</td>
<td>Good patient health, reduce costs</td>
<td>Patient, hospital and its staff</td>
</tr>
<tr>
<td>Satellite image analysis system</td>
<td>Pixels of different intensity, colour</td>
<td>Print a classification of a scene</td>
<td>Accurate classification</td>
<td>Pictures from orbiting satellite</td>
</tr>
<tr>
<td>Refinery controller</td>
<td>Temperature, pressure measurement</td>
<td>Open, close valves; regulate temperature</td>
<td>Maximise purity, optimise yield, wellbeing</td>
<td>Refinery</td>
</tr>
<tr>
<td>Interactive English tutor</td>
<td>Words typed</td>
<td>Print exercise, recommendations, evaluations</td>
<td>Improvement of a student’s marks in exams</td>
<td>Number of students</td>
</tr>
</tbody>
</table>
5. RESEARCH FRAMEWORK

The literature review revealed the four independent variables (Figure 2) that are applied to the dependent variable – cyber security. The impact of such AI technologies on the dependent variable will be further studied.

To comprehensively understand the application of independent variables on dependent variable, appropriate studies will be engaged based on secondary research, and then qualitative analysis performed.

Figure 2. Research framework showing one dependent variable, four independent variables, problems, objectives and outcome, leading to the ‘discussion, analysis and outcome’

6. DISCUSSION, ANALYSIS AND OUTCOME

This paper aims to improve the current understanding of the application of various AI technologies in cyber security and how these emerging technologies may change the way cyber risks are handled.
by companies today. An in-depth analysis shall lead to concrete outcomes of the influence of various constructs on the main research topic.

6.1 Expert Systems – Analysis and Outcome

As expert systems are the leading and most known applicable AI technology in combating against cyberattacks, they are considered the primary tool to safeguard a company’s network.

An expert system consists of a knowledge base in which expert opinion or knowledge is saved regarding a specific application area. Additionally, it implements an inference engine for leading answers considering current information and enhanced knowledge regarding a situation. An expert system shell consists of a vacant knowledge base and inference engine before the loading of its relevant knowledge. To enter information or data based on facts in the knowledge base software, it must support the expert system shell, and it can be loaded with programs for client support, in addition to varied projects that may be used as a part of hybrid expert system (Anwar & Hassan, 2017).

Figure 3 below demonstrates the concept of an expert system at a very basic level:

Figure 3. The expert system concept (Nicolau, Augusto, & Schirru, 2017)

![Expert System Diagram]

Expert systems are categorised as per the view that knowledge is showcased in its knowledge base. The data or knowledge can be shown or presented in many forms, which are usually rules but can also be logic trees and logical frameworks. If the knowledge base is rule-based, its data or knowledge is coded as IF-THEN rules. Otherwise, the same data could be represented via a logic tree model. Hence, the knowledge base model ought to be selected based on the closest illustration to the actual issue or in the most expressive way (Nicolau, Augusto, & Schirru, 2017, June).

The output is only as good as the input. With this in mind, to ensure that knowledge base includes quality data, it is crucial to ensure or implement a quality assurance and control process to verify the accuracy and relevance of data before it goes in the knowledge base. Various quality-oriented models, such as TQM, Deming Cycle and Kaizen, could be applied to improve the quality of data in a knowledge base. Michael Porter’s cost leadership strategy is apt as with a smaller number of attacks, the unit cost of products or services of a company would be minimised, leading to a better quality.

6.2 Neural Networks – Analysis and Outcome

ANNs are webs of connected processing neurons. Each neuron receives a set of mathematical data input from different areas and based on this data, an outcome or output is formed. The outcome is applied to the situation, otherwise it is advanced as input to further network neurons (Demertzis, Iliadis, Avramidis, & El-Kassaby, 2017).
The figure below demonstrates that ANNs have three layers, namely an input layer, hidden layer, and output layer. Figure 4 is a classic example of shallow learning (SL) where there is only one hidden layer. Consecutively, if the neural network model has multiple hidden layers, then it is called Deep Learning (DL).

Figure 4. Basic architecture of a typical multilayer feed-forward ANN (Demertzis, Iliadis, Avramidis, & El-Kassaby, 2017)

All deep learning algorithms (also called a set of rules or coded mathematical formulae) are set upon Deep Neural Networks (DNN), which are huge neural networks prearranged in multiple levels meant for representation learning on their own. Deep learning is recognised to outdo shallow learning in a few application areas, for example, computer vision. This may not always be the case for cyber security in which a few better configured SL algorithms might succeed despite the rarity of DL applications regarding SL procedures in this field (Apruzzese, Colajanni, Ferretti, Guido, & Marchetti, 2018).

Thus, both shallow learning and deep learning neural networks proactively have a significant impact on managing cyber security today, though they may have a separate application. However, each one is important and plays a different role in detecting network anomalies. Various shallow and deep learning models can be implemented as per the requirement in managing Cyber threats, including but not limited to Supervised SL Algorithms such as Random Forest, Support Vector Machines and Logistic Regression and Unsupervised SL Algorithms such as Clustering and Association. Similarly, Supervised DL Algorithms include fully connected feedforward deep neural networks (FNNs), convolutional feedforward deep neural networks (CNNs) and recurrent deep neural networks (RNNs), whereas unsupervised dl algorithms include deep belief networks (DBNs) and stacked autoencoders (SAEs). Since all these models fall under a cybernetics model, a cost leadership strategy could be applied here.

6.3 Intelligent Agent – Analysis and Outcome

Intelligent agents are smart agents created due to their use in cyber security as they are meant for such activities considering their communication and language understanding capability, mobility, rationality, and adaptability. Due to these traits, they are apt to fight against DDoS since the websites or networks become inaccessible due to a large amount of traffic flooding.

Intelligent agents need to go a step further and necessarily have knowledge of previous experiences and the ability to create goals, value specific outcomes and be mindful of their environment. Undoubtedly, when compared with a calculator, they have a greater spectrum of intelligence. They
function in a far more intelligent manner, and systems like natural speech and data processing should be considered AI systems (Shankar, 2017).

Agents have two primary functional modules helping them detect anomalies in the network. Primarily, an agent can learn from the cyberattacks and remember rules based on the previous detections of a basic data set. It notices when data that is incoming breaks these rules. Secondarily, whenever an alert rule is activated, the agent will do an extensive information analysis to realise and understand the actual situation and perform as per the analysis (Wang, & Govindarasu, 2016).

During the DDoS outbreak, the agent will primarily investigate the high incoming volumes toward the application server being targeted. When the volume of requests sent to the server is more than what is predefined, an attack is sensed immediately (Bawany, & Shamsi, 2019).

Based on their perceived intelligence and capability, some of the intelligent agent models are learning agent, utility-based agent, goal-based agent and simple reflex agents, among others. Hence, the cybernetics model could be applied in this case under Michael Porter’s cost leadership strategy.

Figure 5. Utility-based agent model (Source: Wikipedia)

6.4 Artificial Immune System – Analysis and Outcome

The negative selection algorithm is a generation algorithm used to create accurate and efficient detectors that distinguish between self and non-self. The components needed in the negative selection algorithm are the threshold, number of detectors, and the number of features. In the human immune system, an antibody is used by the immune system to neutralise pathogens like viruses and bacteria. Similarly, the Artificial Immune System (AIS) detects its version of an antigen, the intrusion, by the negative selection algorithm. The threshold is what determines an intrusion. The artificial body uses the sensory attribute of the negative selection algorithm whereby a specified number of detectors tripped will cause the data record to be classified as an intrusion (Cooper, A. 2017).
An artificial immune system gets its inspiration from the natural immune system which works competently for sending an unknown anomaly in a network. There are two layers of defences, the innate system and adaptive system, which are deployed in this suggested framework where the innate system copies the natural innate immune system to create the opening line of security or defence. The adaptive system mimics the adaptive immune system by integrating the T-cell and B-cell defensive mechanisms. The outcome displays that the suggested framework works competently for sensing an invasion or anomaly after introducing malicious occurrences to the network system (Dutt, Borah, & Maitra, 2016). AIS makes use of the immune system features, such as feature extraction, recognition of patterns and learning & ability, to logically memorise to change itself with experience and endlessly progress to accomplish an advanced level of precision (Chen, Chang, & Wu, 2016).

From the above, it is understood that AIS is an intelligent rule-based learning system stimulated from the human immune system. Learning and memory are the main characteristics of the immune system that are used for problem-solving in preventing cyberattacks, with negative selection algorithm (typically used for classification and pattern recognition problem) being the model that has been researched and applied under this AI technology.

7. SCOPE FOR FUTURE RESEARCH

This paper observed how some of the AI technologies, such as expert systems, artificial neural networks, intelligent agents and artificial immune systems, are applied to cyber security, indicating their use cases and clear applications benefitting companies or users. However, a myriad of other technologies within and beyond AI, including but not limited to fuzzy logic control systems, search and optimization, and probabilistic methods for uncertain reasoning, can be explored and applied in the cyber security field or industry. AI application and impact could be studied further in terms of the benefits or limitations this technology could bring. It indeed opens the door to a much broader field of application for further study. Also, there is very little information available on AI handling computer virus attacks, which may be largely due to human intervention. Nevertheless, this remains an area requiring more research.
8. CONCLUSION

Artificial intelligence and cyber security are complementary in terms of their application and utilisation. Knowledge acquisition is the most critical part of using expert systems against cyber threats. Quality, completeness and accuracy can bring about a considerable positive change in the way threats are currently handled to improve this precarious area. In expert systems development, a decent solution is based on sound knowledge representation. For expert systems applications, the starting choice of a representation technique is specifically of prime importance (Muhammad, Garba, Oye, & Wajiga, 2018).

It is understood that neural networks, such as shallow learning and deep learning, have taken the cyber security solution one step ahead in the game. Such networks have become advanced technologies to fight cyber threats. Deep learning neural networks are known specifically to learn patterns from the data autonomously as only the input data is provided. In this unsupervised learning, the data to be discovered is known as unlabeled data classified as Self Organization Maps (SOMs) and Adaptive Resonance Theory (ART). A hybrid artificial neural network method is the most appropriate intrusion discovery system regarding detection rates, false positives, false negatives, and cost and timesaving (Hodo, Bellekens, Hamilton, Tachtatzis, & Atkinson, 2017).

Intelligent agents capable of communication and language understanding are apt to act as cyber police in order to fight against DDoS. Multi-agent systems are distributed and can learn on their own from historical anomalies. Their design philosophies and functioning benefits (such as being autonomous and collaboration) permit the producing of real-time and distributed DDoS discovery or detection and systems that can simultaneously discover and bring down various sources or attacks. During a DDoS attack, multiple mobile agents can be posted to the users that are impacted to examine and lessen the effect of network anomalies (Osei, 2018).

As per our research, artificial immune systems (AIS) leverage the human immune system (HIS) by using its own version of an antigen to detect unknown intrusions by using a negative selection algorithm (NSA), which has been discussed in this paper. It is believed that by mimicking HIS, AIS can fight against cyber threats similarly, that is, using its natural form, which in this case, means using NSA. The NSA algorithm based on AIS is utilised to bring down system training time by keeping the correctness in detection. The NSA algorithm is based on self-set (normal) and nonself-set (anomaly) for behaviour discovery. The primary purpose of this arrangement is to identify normal and anomaly processes (Xu et al., 2019). It has a set of self-patterns for recognising anomalies called detectors (Hosseini & Seilani, 2019).

To conclude, Artificial Intelligence (AI) technologies play a significant and critical role in managing precarious tasks related to cyber security. Considering the advancement of attackers, AI could play a crucial role in the coming years. However, one needs to be wary of the limitation this technology could bring, especially in unsupervised learning, where the system is autonomous. This also brings a need for establishing a regulatory environment around AI technologies to ensure adherence to a controlled set of processes and procedures.
REFERENCES


